Abstract

Multilabel classification has attracted much interest in recent times due to the wide applicability of the problem and the challenges involved in learning a classifier for multilabeled data. A crucial aspect of multilabel classification is to discover the structure and order of correlations among labels and their effect on the quality of the classifier. In this work, we propose a structural Support Vector Machine (structural SVM) based framework which enables us to systematically investigate the importance of label correlations in multi-label classification. The proposed framework is very flexible and provides a unified approach to handle multiple correlation orders and structures in an adaptive manner and helps to effectively assess the importance of label correlations in improving the generalization performance. We perform extensive empirical evaluation on several datasets from different domains and present results on various performance metrics. Our experiments provide for the first time, interesting insights into the following questions: a) Are label correlations always beneficial in multilabel classification? b) What effect do label correlations have on multiple performance metrics typically used in multilabel classification? c) Is label correlation order significant and if so, what would be the favorable correlation order for a given dataset and a given performance metric? and d) Can we make useful suggestions on the label correlation structure?